REMARKS

The above amendments to the above-captioned application along with the following remarks are being submitted as a full and complete response to the Office Action dated November 27, 2002 (U.S. Patent Office Paper No. 6). In view of the above amendments and the following remarks, the Examiner is respectfully requested to give due reconsideration to this application, to indicate the allowability of the claims, and to pass this case to issue.

Claims 1, 2, 3, 4, 6, and 11 are currently pending in this application. As outlined above, claims 1, 2, 3, 4 and 11 are being amended to correct formal errors and to more particularly point out and distinctly claim the subject invention. Further, the drawings and abstract are being amended to correct formal errors and to better disclose and describe the features of the present invention as claimed. Applicant hereby submits that no new matter is being introduced into the application through the submission of this response.

Formal Objections or Rejections

Claims 1, 3, 4 and 11 are objected to due to several minor informalities. Applicants have amended the claims and believe that the above mentioned claims, in their amended format, comply with the requirements of 37 C.F.R. 1.75(i) and with the suggestions of MPEP 608.01(m). Applicants respectfully ask the examiner to withdraw the objections.

Claim Rejections under 35 U.S.C. § 112

Claims 2, 3, 4, and 11 were rejected under 35 U.S.C. §112, second paragraph, as being indefinite, in particular, for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Applicants have amended the claims and believe that the amendments to the claims have cured the informalities that were indicated by the Examiner.

In particular, claim 2 has been amended to recite "comparing the expression patterns of two different genes to determine whether are identical or not". In its amended form, claim 2 recites a method step "comparing".

Applicants agree with the Examiner's interpretation of the claim to be reciting "comparison of a reference value to a single expression pattern from each of two different genes to determine whether the expression patterns of the two different genes are identical."

Applicants rewrote claims 3 and 4 to accommodate the Examiner's suggestion regarding claiming the intended method steps in an active, positive language. Applicants believe that by amending the claims 3 and 4 respectively, the informalities regarding these claims have been cured. Applicants respectfully ask the Examiner to withdraw the rejection.

Prior Art Rejections

Claims 1, 2, 6 and 11 were rejected under 35 U.S.C. §102(b) as allegedly being anticipated by Perou *et al.* (PNAS 8/1999, vol. 96, pp. 9212 – 9217), (further, the Perou reference). Applicants respectfully traverse the rejection.

The Perou reference discloses on page 9212, column 2, starting at line 39, that: "Gene-clustering analysis was performed as described in Ref. 12. For the cluster diagrams presented in FIG 1 and 2, the input parameters were to select the subset of genes that has a R/G ration of 3-fold or higher on at least two or more arrays (FIG. 1) or a R/G ration of 3-fold or higher on at least three or more arrays (FIG. 2). The primary data tables and other materials are available at http://genome-www.stanford.edu/sbcmp/."

The Perou reference also discloses on page 9214 a FIG. 2 that illustrates "Overview of the combined in vitro and breast tissue specimen cluster diagram. A scaled-down representation of the 1,247-gene cluster diagram (see Supplemental FIG. 6 at www.pnas.org for the full cluster diagram with all gene names). The black bars show the positions of the clusters discussed in the text: (A) proliferation-associated, (B) IFN-regulated, (C) B lymphocytes, and (D) Stromal cells."

The Perou reference discloses on page 9215 a FIG. 3 that illustrates "Expanded view of two gene clusters taken from the associated cluster diagram. (A) A portion of the proliferation-associated cluster. The numbers below each breast tumor's column show the percentage of carcinoma cells in each specimen that stained positive for the Ki-67 antigen. (B) Expanded view of the IFN-regulated gene cluster. The letters below each breast tumor's columns identify the STAT1 staining pattern seen, with O representing no STAT1 staining (BC17 and FIG> 4A), W representing weak STAT1 staining, S representing strong staining (BC23 and

FIG> 4C), P representing peripheral tumor cell nest staining (BC14 and FIG. 4E), and L representing staining or lymphocytes/histiocytes only (BC16 and Fig. 4G)."

The Applicants' interpretation of the above citations from the reference and about the content of the reference as a whole, is that the Perou reference discloses in FIG. 2 the result of clustering the gene expression patterns and an expanded representation of the proliferation-associated gene cluster and the INF-associated gene cluster; see FIG. 2(A), 2(B), and 3.

Regarding FIG. 3, the illustration outline reads "Expanded view of two gene clusters taken from the associated cluster diagram."

In the Office communication, on page 5, the Examiner alleged that "Perou further teaches comparison of data form expression in different tissue states to a reference value, and "repeating" clustering such that clusters "move" along an axis by comparison to changing, or different reference values (e.g. figure 3 shows different (moved) clusters based on comparison to IFN-regulated reference genes and comparison to proliferated-associated reference genes), thus claims 1 and 11 are anticipated."

On the contrary with the Examiner's allegation, Applicants submit that claims 1 and 11 comprise the steps off "clustering the expression pattern data within the designated segment along the second axis based on a predetermined reference value; repeating clustering within the same cluster in a forward or reverse direction along the second axis while changing the reference value, and displaying the results according to a predetermined display format.". In the claims the term "reference value" is used in connection with the step "clustering" and "repeating clustering", not with the "comparison of data", as suggested by the Examiner.

In light of the above remarks, Applicants respectfully ask the Examiner to withdraw the rejection regarding claims 1 and 11. The Perou reference does not anticipate the claimed matter of claims 1 and 11 because it does not disclose every feature of the claimed matter.

Claims 2 and 6 depend from and add features to an allowable claim, therefore they are also allowable. Applicants respectfully ask the Examiner to withdraw the rejection regarding claims 2 and 6.

Claims 1, 2, 3, 4, 6 and 11 stand rejected under 35 U.S.C. §102(e) as allegedly being anticipated by Zheng *et al.* (U.S. Patent No. 6,263,2870, the Zheng reference). Applicants respectfully disagree.

In the Office Communication, on page 6, the Examiner points out that "Zheng teaches a method and apparatus for displaying gene expression patterns of genes related to disease (tissue) states characterized by differences in expression patterns wherein gene expression data from a plurality of genes is compared on a two-dimensional grid comprising gene and expression data, and expression patterns are compared to a reference (col. 5, line 16 – col. 6, line 19 and FIG. 13)." The Examiner specifically points out that "Zheng specifically teaches comparison of expression levels in different tissues or disease states (col. 7, lines 14-26), thus claims 1, 6, and 11 are anticipated."

The Applicants want to respectfully draw the Examiner's attention onto the fact that the Zheng reference discloses that sets of digits corresponding to the relative expression levels ("grid representation") are used as reference value. See the adjoined paragraph from the Zheng reference, specifically from col. 6, line 48 through col. 5, line 9:

"In a preferred embodiment, the GED is assigned a grid representation using the following methodology: (1) each GED time curve [E.sub.i,1, E.sub.i,2, . . . , E.sub.i,M] is coarse-grain averaged to [<E.sub.i,1 >, <E.sub.i,2 >, . . . , <E.sub.i,m >], where <E.sub.i,j > denotes an arithmetic average over the time points within time stage j; (2) <E.sub.i,j > is rounded to the nearest integer of (<E.sub.i,j >/.DELTA.E), denoted by E*.sub.i,j ; and (3) the grid representation of the original time curve i with m stages and n levels is represented by [E*.sub.i,1, E*.sub.i,2, . . . , E*.sub.i,m ; n].

Grid representations preferably are used to provide a simple naming mechanism for clustering the GED. For example, one may describe a differential gene expression curve with five time points as follows: "the expression is initially up-regulated, then becomes significantly up-regulated and stays there until the fourth time point, when the up-regulation becomes moderate, and finally returns to normal at the end." With the grid clustering, such a curve can be described in a grid representation as [1, 2, 2, 1, 0; 5]. The last digit "5" in this grid representation indicates that the grid has five relative expression levels: -2, -1, 0, 1, and 2, corresponding to, in a preferred embodiment, significantly down-regulated, down-regulated, normal, up-regulated and significantly up-regulated expression. One skilled in the art understands that the number of time

points and relative expression levels chosen for grid representation naming is not limited in any fashion and that the systems of the invention are fully adaptable in this regard. Other benefits of this grid representation naming mechanism include the ability to search, sort, and present data, as well as perform arithmetic operations within the context of the present invention, as described supra. ".

In the Zheng reference, genes with the same grid representation are grouped as representing an expression pattern and grid clustered. See the adjoined paragraph from the Zheng reference, specifically from col. 7, line 54 through col. 8, line 50:

"1. Grid Clustering

Unlike the sequence-related clustering based on the established sequence and function correlation, the clustering of time curves to identify the functional correlation of genes is inherently uncertain. This is because genes with similar time curves are not necessarily functionally related, and functionally related genes may exhibit very different time curves. Indeed, FIG. 8 provides a comparison of two representative genes whose functional correlation involves a scale change, a time delay, and a vertical flip, respectively.

Clustering analysis is an important tool, since it helps in reducing the complex pattern of thousands of time curves into a smaller set of representative clusters. The systems of the present invention allow one to cluster and view the curves in many different ways. This preferably maximizes the chance of capturing the functional correlation of genes. Indeed, the grid and .sigma.-.tau. clustering algorithms of the systems of the present invention are preferably used for clustering time curves and thus assessing the functional correlation of genes.

In a preferred embodiment of the present invention, GED assigned a grid representation that may be grid clustered. This aspect of the present invention transforms the process of clustering many curves into a smaller number of representative clusters into a process of coarse-grain averaging the curves onto a two-dimensional grid. This averaging process is fast (O(N)), hierarchical and unambiguous. Grid clustering may be accomplished by binding curves onto a two dimensional grid with m (0<m.ltoreq.M) time stages and n (n>1) expression levels.

Each curve belongs to a cluster defined by the grid representation of the curve. With the exception of the last time stage if M/m gives a remainder, each time stage contains M/m time points. For example, for M=10, m=3, the 10 time points are partitioned into 3 time stages as

(1, 2, 3, 4), (5, 6, 7, 8) and (9, 10). Each discrete expression level covers an interval of the continuous expression value:

.DELTA.E=(max{E.sub.i,j }-min {E.sub.i,j })/(n-1)

For {E.sub.i,j } normalized to [0, 1], the length of each interval is 1/(n-1), and the discrete expression levels are 0, 1, . . . , n-1. For {E.sub.i,j } normalized to [-1, 1], it is preferable to choose an odd number for n so that the negative, 0 and positive levels can be evenly represented. For example, for n=2k+1, where k is a positive integer, the length of each interval is 1/k, and the discrete expression levels are -k, . . . , -1, 0, 1, . . . , k.

Each time curve preferably is associated with a unique cluster. The geometric shape of a cluster preferably is explicitly represented by the cluster's grid representation name.

FIG. 5 is a flowchart of another preferred embodiment of the systems of the present invention that shows the clustering of processed GED through Grid Clustering 360. If Grid Clustering 290 is desired, the systems of the present invention provide for taking GED from the Processed GED Store 350 and Grid Cluster 360 the processed GED. Grid clustered GED may then be presented graphically 370 for the user to see. Once displayed, the user may then determine if the grid size is too coarse 380. If the grid size is too coarse, the user may reduce the grid size by means of the keyboard 140, mouse 150, or other such hardware/software allowing input of data to the computer system 100. See Section P., infra. If the grid size is appropriate, the user may then search and manipulate the data as shown in the flowchart of FIG. 7. "

In the Zheng reference the grid clustering using the grid representation also allows detecting the same gene expression curves (see Fig. 8C). More, it also allows detecting mutually reversal gene expression curves (as shown in Fig. 8D). For further details, please refer to the section of the Zheng reference found between col. 8, line 51 and col. 10, line 2.

In the '287 reference, Zheng observes more than one gene expression pattern: different tissues or different disease models. Specifically, see the adjoined paragraph, from the Zheng reference, col. 7, lines 14 through 27:

"Another useful feature of this naming mechanism of the grid representation is that the difference between two time curves of the same gene (e.g., the expression level in different tissues or different disease models) can be conveniently expressed as the difference between the two individual cluster names. For example, let [E*.sub.i,1, E*.sub.i,2, . . . , E*.sub.i,m ; n] and [E*'.sub.i,1, E*'.sub.i,2, . . . , E*'.sub.i,m ; n] denote the two time curves.

Their quantitative difference can be measured by DELTA.E*.sub.i.ident.E*.sub.i - E*'.sub.i.ident.[E*.sub.i,1 -E*'.sub.i,1, E*.sub.1,2 -E*'.sub.i,2, . . . , E*.sub.i,m -E*'.sub.i,m ; n]. This compact form is convenient, for example, in searching for tissue and disease specific expression patterns within the context of the present invention."

In the present invention, Applicants are clustering the data of more than one expression patterns using a reference value. But, the present invention distinguishes itself by clustering within the same cluster while changing the reference value. This is a feature not contemplated by the Zheng reference. Therefore, at least for this reason, the Zheng reference does not anticipate the present invention.

Furthermore, in the Zheng reference the "grid representation" corresponds to the relative expression levels. In the pending application, the reference value determines whether two expression patterns of different genes are identical or not.

Based on the remarks expressed above, Applicants submit that the Zheng reference does not disclose all the features of claims 1, 6, and 11. Applicants respectfully ask the Examiner to withdraw the rejection regarding the above referenced claims.

Further, the Examiner has alleged that claims 2, 3, and 4 are anticipated by the Zheng reference (Col. 15 lines 49 through 67) because Zheng's "method and system can be used to determine if expression patterns from different genes are identical, converge or diverge."

The above citation shows that a user can type grid representations, such as [0,0,0,0,0;5] and [0,0,0,3,3;5] into the text field, to search for the correspondent gene expression patterns. In this case, the user has to repeat the typing and searching to see if these expression patterns are identical, converge or diverge.

On the other hand, the present invention allows to automatically determine and display such information. In particular, in Claim 3, "displaying two or more different genes according to the predetermined display format, whereby said two or more different genes have the same expression pattern at the beginning of said experiment case but change to different expression patterns within the segment along the second axis." Respectively in claim 4 "displaying two or more different genes according to the predetermined display format, whereby said two or more different genes have different expression patterns at the beginning of said experiment case but change to the same expression pattern within the segment along the second axis."

Based on the above remarks, the Applicants respectfully traverse the Examiner's anticipation rejection relative to claims 2, 3, and 4. The Examiner is respectfully asked to withdraw the rejection.

The present invention and the cited references are different in their methods for clustering and displaying. Perou is drawn to clustering different expression patterns and displaying part of a certain cluster in an expanded view. Zheng is drawn to classifying expression patterns according to the grid representation and displaying the expression patterns from the genes having the same grid representation.

The present invention, however, is drawn to repeating clustering into a small cluster size while changing the reference value for the clustering. The resultant expression patterns are automatically displayed, such as shown in FIGs.2 and 3 of the present invention. Such automatic displaying or other methods for automatic displaying are not anticipated by Perou or Zheng.

It is an advantage of the present invention that with such automatic displaying, a user can address to more detailed representations of the gene expression patterns. For example, the user can search for a pair of different genes having the same expression pattern at the beginning but growing to have different expression patterns with reference to FIG. 2. Thus, the user can bring forth biological interpretations, e.g. presumption on the interactions between the genes. Perou and Zheng allow a user to study more than one gene having the same expression patterns, but fails to allow studying more complicated expression patterns such as in the above.

CONCLUSION

In view of all the above, Applicants respectfully submit that certain clear and distinct differences as discussed above exist between the present invention as now claimed and the prior art references upon which the rejections in the Office Action rely. These differences are more than sufficient that the present invention as now claimed would not have been anticipated nor rendered obvious given the prior art. Rather, the present invention as a whole is distinguishable, and thereby allowable over the prior art.

Favorable reconsideration of this application as amended is respectfully solicited. Should there be any outstanding issues requiring discussion that would further the prosecution and allowance of the above-captioned application, the Examiner is invited to contact the Applicants' undersigned representative at the address and phone number indicated below.

Respectfully submitted,

Stanley P. Fisher
Registration Number 24,344

Juan Carlos A. Marquez
Registration Number 34,072

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200

February 27, 2003

MARKED-UP COPY OF THE CLAIM AMENDMENTS

1. A method for displaying gene expression patterns of multiple genes whose expressions change according to experiment cases, where a first axis represents the genes and a second axis represents the experiment cases, the method comprising the steps of:

designating a segment along the second axis in the expression pattern data of the multiple genes; [and]

clustering the expression pattern data within the designated segment along the second axis based on a predetermined reference value[,] ;

repeating clustering within the same cluster in a forward or reverse direction along the second axis while changing the reference value, and

displaying the results according to a predetermined display format.

2. A method for displaying gene expression patterns according to claim 1, wherein comparing the [reference value determines whether two] expression patterns of two different genes to determine whether are identical or not.

3. A method for displaying gene expression patterns according to claim 1 or 2, further comprising:

displaying [wherein] two or more different genes [are displayed] according to the predetermined display format,

[where] whereby [they] said two or more different genes have the same expression pattern at the beginning of said experiment case but [become to have] change to different expression patterns within the segment along the second axis.

4. A method for displaying gene expression patterns according to claim 1 or 2, further comprising :

displaying [wherein] two or more different genes [are displayed] according to the predetermined display format,

[where] whereby [they] said two or more different genes have different expression patterns at the beginning of said experiment case but [become to have] change to the same expression pattern within the segment along the second axis.

6. A method for display gene expression patterns according to claim 1, wherein the experiment cases are states of individual's tissue.

11. An apparatus for analyzing gene expression patterns, which acquires, from a database, expression pattern data of multiple genes whose expressions change according to experiment cases, and which visually displays the expression patterns on a screen of a display device, where a first axis represents the genes and a second axis represents the experiment cases, the apparatus comprising:

an inputting means for designating a segment along the second axis in the expression pattern data of the multiple genes obtained from the database[;] , and

an arithmetic unit for clustering the expression pattern data within the designated segment along the second axis based on a predetermined reference value,

that [repeating] repeats clustering within the same cluster in a forward or reverse direction along the second axis while changing the reference value, and

[displaying] displays the results according to a predetermined display format.